49
~~75~~

Claims

What is claimed:

1. An invention substantially as described in the description.

2. An invention substantially as described and illustrated in the description.

3. A process for identifying one or more bi-allelic markers linked to a bi-allelic genetic characteristic gene in a species of creatures, comprising the steps of :

a)choosing two or more bi-allelic covering markers so that a CL-F region is systematically covered by the two or more covering markers, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency;

b)choosing a statistical linkage test based on allelic association for each covering marker;

c)choosing a sample of individuals for each covering marker ;

d)obtaining genotype data/sample allele frequency data for each covering marker and the sample chosen for each covering marker, and obtaining phenotype status data for the genetic characteristic for each individual in the sample chosen for each covering marker;

e)calculating evidence for linkage between each covering marker and the gene using the statistical linkage test based on allelic association chosen for each covering marker and the genotype data/sample allele frequency data for each covering marker and using the phenotype status data for the genetic characteristic for each individual in the sample chosen for each covering marker obtained in d); and

f)identifying those covering markers as linked to the genetic characteristic gene which show evidence for linkage based on the calculations of step e).

24. A process as in claim 3 wherein the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance δ of N or more of the covering markers, wherein δ is equal to about [δ_{CL}, 0.25] or the equivalent thereof, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has

50
~~76~~

been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

3 ~~5~~. A process as in claim ~~4~~ 2 wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length L_{MC} and width W_{MC}, and L_{MC} being less than or equal to about $2\delta_{CL}$, and W_{MC} being less than or equal to 0.5, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprising essentially less than all of the polymorphisms in the region of interest.

6. A process as in claim 5, wherein the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment.

7. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers.

4 ~~8~~. A process as in claim ~~5~~ 3 wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

9. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof.

10. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional

58
~~84~~

32. An apparatus for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location as in claim 31, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

~~33~~ 7. A process for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising:

a)determining information on the presence or absence of each allele of each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, a CL-F region being systematically covered by the two or more bi-allelic covering markers, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency; and

b) transforming the information of step a) into genotype data/sample allele frequency data for each marker of the group.

8 ~~34~~. A process for obtaining genotype data/sample allele frequency data as in claim 7 ~~33~~, wherein the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance δ of N or more of the covering markers, wherein δ is equal to about [δ_{CL}, 0.25] or the equivalent thereof, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

9 ~~35~~. A process for obtaining genotype data/sample allele frequency data as in claim 8 ~~34~~, wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length L_{MC} and width W_{MC}, and L_{MC} being less than or equal to about $2\delta_{CL}$, and W_{MC} being less than or equal to 0.5, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprising essentially less than all of the polymorphisms in the region of interest.

59
85

36. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment.

37. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers.

9
10 38. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

39. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof.

40. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

41. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele

74
~~100~~

137~~8~~. One or more copies of a set of oligonucleotides, the set of oligonucleotides being substantially complementary to a group of two or more bi-allelic covering markers of the same species, wherein the group of covering markers systematically cover a CL-F region, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency.

13
147~~9~~. One or more copies of a set of oligonucleotides as in claim ~~78~~, wherein the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance δ of N or more of the covering markers, wherein δ is equal to about [δ_{CL}, 0.25] or the equivalent thereof, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

13
158~~0~~. One or more copies of a set of oligonucleotides as in claim ~~78~~, wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length L_{MC} and width W_{MC}, and L_{MC} being less than or equal to about $2\delta_{CL}$, and W_{MC} being less than or equal to 0.5, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprising essentially less than all of the polymorphisms in the region of interest.

81. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment.

82. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers.

15
168~~3~~. One or more copies of a set of oligonucleotides as in claim ~~80~~, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

AMENDED SHEET

probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

11. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

12. A process as in claim 10, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

13. A process as in claim 11, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or

52
~~78~~

less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

14. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells of length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

15. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele

frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells of length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

16. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection D meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells of length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

54
~~80~~

17. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection D meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

18. A process as in claim 4, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

criterion c) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers;

criterion d) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5;

criterion e) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof;

criterion f) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion g) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate

chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion h) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent;

and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P

57
~~83~~

is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

19. A process as in claim 18, wherein δ is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

20. A process as in any one of claims 3-19, wherein there is a subgroup of the covering markers, and the markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

21. A process as in claim 20, wherein the process comprises a computer program.

22. A process as in claim 20 wherein L_{MC} is less than or equal to about 250,000 bp or the equivalent thereof, W_{MC} is less than or equal to about 0.15, wherein the species is human being, wherein the same statistical linkage test based on allelic association is chosen for each covering marker in step b).

23. A process as in any one of claim 22, wherein the process comprises a computer program.

24. An apparatus for identifying bi-allelic markers linked to a bi-allelic genetic characteristic gene in a species of creatures, comprising: means to practice each of the steps of a process as in claim 20.

25. An apparatus as in claim 24, wherein the apparatus comprises oligonucleotide technology or mass spectrometry.

26. An apparatus for identifying bi-allelic markers linked to a bi-allelic genetic characteristic gene in a species of creatures, comprising: means to practice each of the steps of a process as in any one of the claims 3-19.

27. An apparatus as in claim 26, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

28. A process for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location, comprising the steps of: a process as in claim 20, further comprising the step f) of: localizing the gene to the CL-F location of one or more markers that show evidence for linkage based on the calculations of step e).

29. A process for localizing a bi-alleic genetic characteristic gene in a species of creatures to a CL-F location, comprising the steps of: a process in any one of the claims 3-19; further comprising: the step f) of: localizing the gene to the CL-F location of one or more markers that show evidence for linkage based on the calculations of step e).

30. A process as in claim 29, wherein the process comprises a computer program.

31. An apparatus for localizing a bi-alleic genetic characteristic gene in a species of creatures to a CL-F location, comprising: means to practice each of the steps of a process as in claim 29.

40
86

frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

42. A process for obtaining genotype data/sample allele frequency data as in claim 40, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

43. A process for obtaining genotype data/sample allele frequency data as in claim 41, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the

61
87

marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

44. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

45. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC}

in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

46. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

47. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a

chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

48. A process for obtaining genotype data/sample allele frequency data as in claim 34, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

criterion c) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers;

criterion d) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele

frequency of each of the markers in the subgroup being close to 0.5;

criterion e) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof;

criterion f) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion g) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion h) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the

65
~~91~~

group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent;

and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4), wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

49. A process for obtaining genotype data/sample allele frequency data as in claim 48, wherein δ is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

~~50~~. A process for obtaining genotype data/sample allele frequency data as in any one of claims 33-49, wherein there is a subgroup of the covering markers, and the markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

51. A process for obtaining genotype data/sample allele frequency data as in claim 50, wherein the process comprises a computer program.

66
92

52. A process for obtaining genotype data/sample allele frequency data as in claim 50 wherein L_{MC} is less than or equal to about 250,000 bp or the equivalent thereof, W_{MC} is less than or equal to about 0.15, wherein the species is human being.

53. A process for obtaining genotype data/sample allele frequency data as in claim 52, wherein the process comprises a computer program.

54. An apparatus for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising: means to practice each of the steps of a process as in any one of the claims 33-49.

55. An apparatus for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising: means to practice each of the steps of a process as in claim 50.

56. An apparatus as in claim 55, wherein the apparatus comprises oligonucleotide technology or mass spectrometry.

57. An apparatus as in claim 54, wherein the apparatus comprises oligonucleotide technology or mass spectrometry.

58. An apparatus as in claim 54, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

59. A use of one or more copies of a set of oligonucleotides to determine genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers for one or more individuals, each individual being a member of the same species, wherein the group of covering markers systematically cover a CL-F region, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency.

60. A use as in claim 59, wherein the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance δ of N or more of the covering markers, wherein δ is equal to about [δ_{CL} 0.25] or the equivalent thereof, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

61. A use as in claim 59, wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal

67
~~93~~

to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length L_{MC} and width W_{MC}, and L_{MC} being less than or equal to about $2\delta_{CL}$, and W_{MC} being less than or equal to 0.5, δ_{CL} is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprising essentially less than all of the polymorphisms in the region of interest.

62. A use as in claim 61, wherein the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment.

63. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers.

64. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

65. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof.

66. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering

marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

67. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

68. A use as in claim 66, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

69. A use as in claim 67, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and

69

wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

70. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

71. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the

same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

72. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

73. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers

are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

74. A use as in claim 60, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

criterion c) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers;

criterion d) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of

the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5;

criterion e) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof;

criterion f) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion g) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion h) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering

markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent;

and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

75. A use as in claim 74, wherein δ is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

76. A use as in any one of claims 59-75, wherein there is a subgroup of the covering markers, and the markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

77. A use as in claim 76 wherein L_{MC} is less than or equal to about 250,000 bp or the equivalent thereof, W_{MC} is less than or equal to about 0.15, wherein the species is human being.

84. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof.

85. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

86. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

87. One or more copies of a set of oligonucleotides as in claim 85, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were

76
102

chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

88. One or more copies of a set of oligonucleotides as in claim 86, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

89. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC}

in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

90. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length L_{MC} and width W_{MC} in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

91. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the

group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

92. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

93. One or more copies of a set of oligonucleotides as in claim 79, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

criterion c) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers;

criterion d) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5;

criterion e) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof;

criterion f) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion g) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance δ by the two or more bi-allelic covering markers;

criterion h) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent;

and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from

81

~~107~~

substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F region is N covered to within a CL-F distance δ by the two or more bi-allelic covering markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

94. One or more copies of a set of oligonucleotides as in claim 93, wherein δ is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

SUB C5

95. One or more copies of a set of oligonucleotides as in any one of claims 78-94, wherein there is a subgroup of the covering markers, and the markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

96. One or more copies of a set of oligonucleotides as in claim 95 wherein L_{MC} is less than or equal to about 250,000 bp or the equivalent thereof, W_{MC} is less than or equal to about 0.15, wherein the species is human being.

ADD C67

add B5

add B6